MIT Sloan School of Management

Working Paper 4459-04
January 2004

WIRELESS GRIDS: APPROACHES, ARCHITECTURES, AND TECHNICAL CHALLENGES

Ashish Agarwal, Douglas O. Norman, Amar Gupta

WIRELESS GRIDS: APPROACHES, ARCHITECTURES, AND TECHNICAL CHALLENGES

Ashish Agarwal (i2 Technologies)
Douglas O. Norman (The MITRE Corporation)
Amar Gupta (MIT)

ABSTRACT

Grid computing and grid topologies are attracting a growing amount of attention. Originating as a concept for sharing computing resources among wired participants, the grid concept is gradually been extended into the wireless world. A Wireless Grid is an augmentation of a wired grid that facilitates the exchange of information and the interaction between heterogeneous wireless devices. While similar to the wired grid in terms of its distributed nature, the requirement for standards and protocols, and the need for adequate Quality of Service; a Wireless Grid has to deal with the added complexities of the limited power of the mobile devices, the limited bandwidth, and the increased dynamic nature of the interactions involved.

Depending on the nature of the interactions among the constituencies served by the wireless grid, various layouts can be envisaged. The ability of these models to address needs at the enterprise, partner, and service levels is contingent upon the efficient resolution of multiple technical challenges of the grid.

1 INTRODUCING “THE WIRELESS GRID”

Foster [1] offers a checklist for recognizing a “grid”. A Grid allows

- Coordination of resources that are not subject to centralized control;
- Use of standard, open, general-purpose protocols and interfaces; and
- Delivery of nontrivial qualities of service.

The emergence of the Wireless Grid meets all these criteria and is fueled by technological advances in two major areas:

***Grid Computing**:* This is envisaged to provide a solution to the challenge of ‘*flexible, secure, and coordinated resource sharing among dynamic collections of individuals, institutions and resources*’ [2]. The ultimate vision of the grid is that of an adaptive network offering secure, inexpensive, and coordinated real-time access to dynamic, heterogeneous resources, potentially traversing geographic, political and cultural boundaries but still able to maintain the desirable characteristics of a simple distributed system, such as stability, transparency, scalability and flexibility. The technologies originally developed for use in a wired environment are now being augmented to operate in wireless situations.

Wireless Technology: Rapid advances, diminishing prices, wide availability, and attractive form factors have caused wireless technologies to permeate the lives of people from all walks of life.

The development of the wireless technologies such as 802.11, GPRS and 3G has extended the reach of wireless services to (literally) the man on the street. With the ubiquity and indispensability of wireless technologies established, these technologies are now making inroads into Grids.

2 KEY CHARACTERISTICS

2.1 Driving Forces

The development of the wireless grid technologies is governed by three driving forces:

New User Interaction Modalities and Form Factors: Traditional applications that can exist on the Wired Grid need to expand their scope by extending the interactions to mobile devices through adapting the user interface to small screens, small keyboards, and other I/O modalities such as speech. The mobile access interface needs to address the issue of connectivity of mobile devices.

Limited Computing Resources: Wireless applications need to share the resources and provide access to additional computational resources to mitigate the constraints imposed by limited storage, computational capability, and power of mobile devices.

Additional New Supporting Infrastructure Elements: New applications, especially ones involving dynamic and unforeseen events, need to be addressed through the rapid provisioning of major amounts of computational and communications bandwidths. For example, the occurrence of an urban catastrophe could trigger a dynamic adaptive wireless network to alert people to organize remedial actions in a coordinated fashion, and to provide better control of available resources and personnel.

2.2 Grid Resources

A Wireless Grid must provide a virtual pool of computational and communications resources to consumers at attractive prices. Various grid resources are described below:

Computing Power: Wireless devices possess limited computation power. Wireless grids can overcome this limitation by distributing the computational tasks across multiple power-constrained devices. But this raises the need for establishing appropriate collaborative processes between these geographically distributed tasks.

Storage Capacity: Wireless devices possess limited storage capability. Grids can overcome this limitation by distributing the data storage over multiple devices. Data can be recombined into a single entity and then made available to the users. However, this creates the need to enable data access and update to occur simultaneously and to avoid contention through the application of advanced synchronization techniques.

Communications Bandwidth: Wireless grids can harness the power of wireless technology to allow remote access. At the same time, the grid infrastructure should be robust enough to ensure high Quality of Service (QoS).

Multiplicity of Applications: Wireless Grids should allow the users ubiquitous access to a wide variety of applications. However, one needs to overcome the need to install these applications on separate mobile devices.

3 GRID LAYOUT

Drawing upon the paradigm of the wired grids [3,4,5], various layouts of the wireless grids are possible. The classification schemes can be based on the architecture or on the function of the grids.

3.1 Classification by Architecture

One way to characterize the architecture of the wireless grid is by the degree of heterogeneity of the actual devices and the level of control exercised by those who own and administer the devices (Fig. 1). It can vary from a simple network of homogeneous devices bound by a single set of policies and rules to a complex network of heterogeneous devices spread across multiple organizational, political and geographical boundaries, as categorized below:

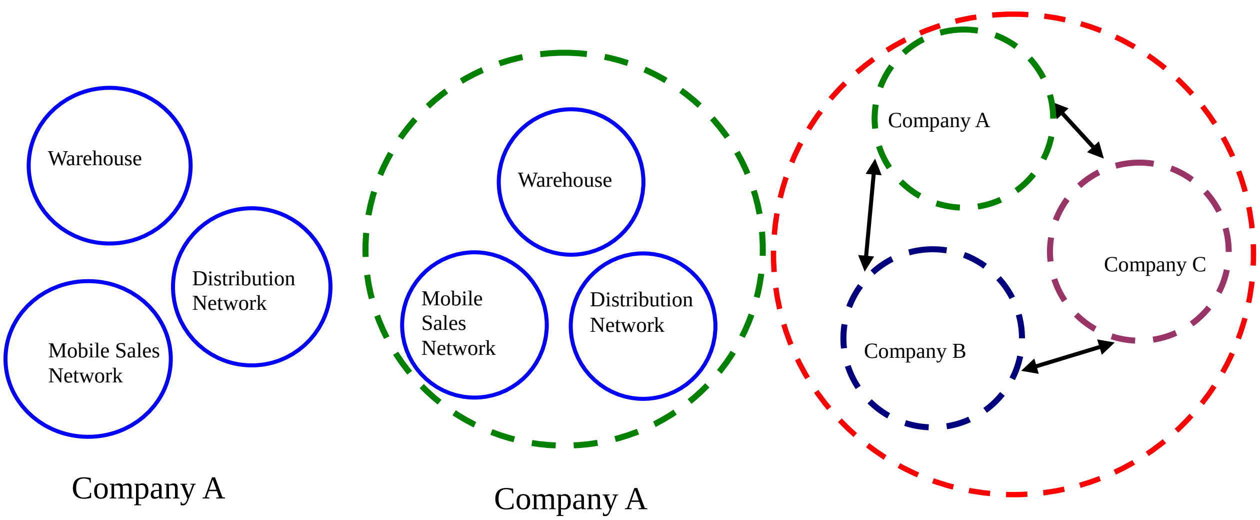
Local Cluster or Homogeneous Wireless Grid: This simplest form involves a local collection of identical or similar wireless devices that share the same hardware architecture and the same operating systems. Because of the homogeneity of the end systems, the integration of these devices into the wireless grid, as well as the consequent sharing of resources, becomes a much easier task. Today, this type of organization is more likely to be found in a single division of an organization where one single administrative body exercises control over all the devices. An example would be a network of mobile handheld devices for coordinating medical personnel in the hospital. It remains to be seen whether market forces will result in convergence of hardware (virtual or real) and software and the emergence of a dominant design, which can exploit resource-sharing strategies that are more intimately bound to the device.

Wireless Intra-Grids: An intra-grid encompasses wireless devices that belong to multiple divisions or communities within an actual organization (AO). The divisions may be located in different geographies and maybe governed by a separate set of policies, but there exists a level of trust and oversight so that “ground truth” may be known with respect to identity and characteristics. AOs are the point where resolution can occur between the virtual presence of a wireless entity and its actual name and location. AOs also tend to be persistent in time, and become the point of composition among other AOs. An example of an intra-grid would be a wireless grid that simultaneously supports the mobile sales force of a company and the networks of wireless sensors used by the manufacturing division for tracking inventory.

Inter-Grid*:* An inter-grid encompasses multiple AOs and transcends greater amounts of geographical, organizational, and other types of differences, such as ones related to intellectual property rights and national laws. Multiple AOs may come together to form Virtual

Organizations (VOs) where they can collaborate and share resources such as information, knowledge, and even market access to exploit fast-changing market opportunities. The relationship can be long or short term [4]. Resource management and policy integration (security, authentication and data management tasks) attain greater complexity due to the scalability requirements. To move beyond mere *ad hoc* composition of AOs, a (potentially) universally accepted method for the composition of declarative policies must be proposed and accepted. Additionally, there must be commonly accepted semantics for the expression of policy. This could be rendered through on-line-available references to expressions and ontologies that would be used by policy-composition and enforcement processes. Without such semantic agreements, and without mechanisms of composition, arbitrary membership in a protected (and trusted) inter-grid set of relations would be impossible. An example of an inter-grid interaction would be a scenario involving an American tourist visiting Japan and trying to conduct a local e-commerce transaction using his/her cell phone. The transaction would involve a handshake between the traveler's cell phone service provider, traveler's credit card company, the Japanese wireless service provider and the e-commerce vendor.

Local Clusters **Intra-Grid** **Inter-Grid**



Local wireless cluster deployed on a departmental/divisional basis within a company.

Merging Cluster Grids into an Intra Grid that is within a company.

Merging Intra Grids of many companies into an Inter Grid that spans multiple organizations

Figure 1: A Simplified Depiction of the 3-Tier Wireless Grid Architecture [4]

3.2 Classification by Usage Pattern

Wireless grids can be classified by usage patterns as summarized in Table 1.

Computational Grid: In a computational grid, the need for creating the wireless grid is driven primarily by the need to borrow computational resources from others. This arises, in part, because of the power constraints on mobile devices, which in turn limits their computational capability. The computational grid may be cooperative or parasitic [6]. One example is a wireless sensor network used to monitor conditions for predicting natural calamities like earthquakes or volcanoes.

Data Grid: In this case, the need for creating the wireless grid is dictated primarily by the need to provide shared and secure access to distributed data. Since data can be presented in various contexts on various systems, reconciling the underlying semantics continues to challenge evolving technology. One example involves an urgent search for donors with a rare blood type. A hospital would issue a query to the medical history databases in the region through its mobile network. The mobile service providers will notify potential donors through the alert messages transmitted to their respective mobile devices, and the resulting responses would be processed and reconciled.

Utility Grid: Here the motivation for the wireless grid is derived from the need to provide ubiquitous access to specialized pieces of software and hardware. Users can request resources when needed (on-demand) and only be charged for the amount being used. This model can subsume both Computational and Data grids. For example, users might tap Wireless Utility grids for information such as the traffic conditions and routing, and for making instantaneous transactions related to commercial products and services.

Table 1: Wireless Grid Usage Patterns

Grid Type	Possible Architecture	Mainly Provides
Computational	Cluster, Intra, Inter	Computational Power
Data	Cluster, Intra, Inter	Data Access and Storage
Utility	Intra, Inter	On-demand Access to All Kinds of Resources

4 GRID TOPOLOGY

A number of researchers have evaluated the topology and configuration of mobile networks

[7,8,9]. However, these ad hoc systems are standalone in nature. We believe that the commercial grids will possess some access to the wired Internet infrastructure and thereby follow a hybrid model (Fig 2). It will consist of Mobile Ad-hoc Networks[1] (MANET) type systems with multiple-hop paths between mobile nodes and access points to the wired network.

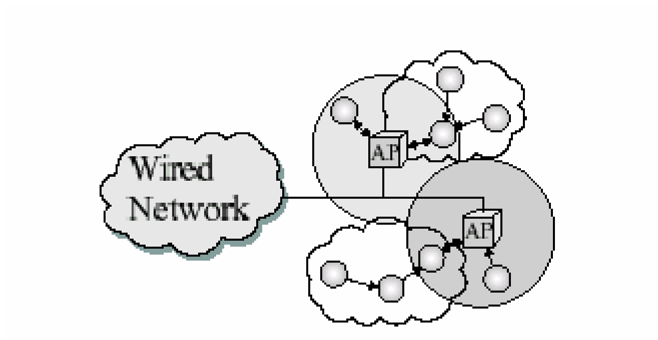


Figure 2. A Hybrid Wireless Network

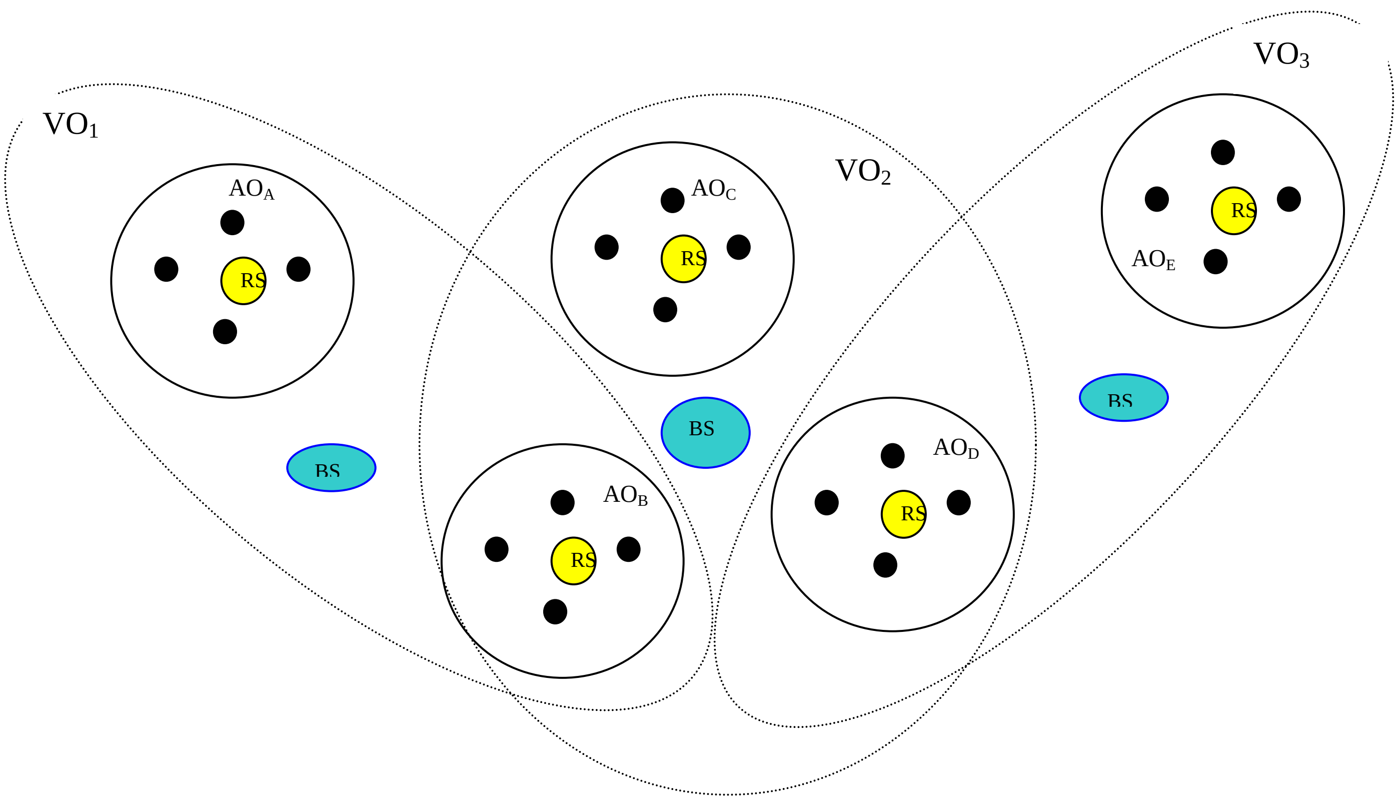


Figure 3. Wireless Grid spanning multiple Virtual Organizations

[1] http://www.ietf.org/html.charters/manet-charter.html

At a high level, one needs to support the critical role of the management and composition of subnets and arbitrary collections of wireless members. There must be a Root Station (RS) present in some form as well as a Base Station (BS). The RS maintains cognizance over a set of wireless devices and serves as the final mapping of logical to physical devices. The BS manages and enforces policy within and among groups. A grid layout can include a root station for a community or an actual organization (AO) of wireless nodes (Fig. 3). A RS will maintain up-to-date information about its own network and the associated nodes as well as serve as the gateway to the wired network. Multiple organizations may come together to form a virtual organization (VO). An AO can belong to multiple VOs. A base station (BS) can be envisaged for a VO. A BS will maintain information about networks for various organizations and the associated root stations. For a homogeneous grid, the same server can perform both the RS and BS functions. In case of an inter-grid that can span multiple virtual organizations, several BSs are needed to coordinate to maintain the inter-grid information. Redundancy can be maintained by having secondary servers to perform the RS and BS functions. Both RS and BS should not be resource-constrained devices. Instead, the RS and the BS could each be a simple PC, a workstation, or a server equipped with an appropriate interface to communicate with the edge nodes such as sensor nodes and other mobile nodes.

5 TECHNICAL CHALLENGES

Among the challenges related to wireless grids, one must overcome the following set of initial technical challenges:

Dynamic Configurability: Wireless grids are characterized by changing topology due to the mobile nature of the grid components. Grids should provide self- configuring and self-administering capability to allow these dynamic changes for all possible grid layouts. This includes configuring the addresses, providing name to address resolution for the grid components, and maintaining the state information. Grids should address the issues related to weak transmission signals, message losses, and crashing of the power constrained nodes.

Routing Plasticity: Efficient routing protocols are required to address the power limitation of the end devices along with the consideration for stable wireless connectivity, route optimization and efficient use of the limited bandwidth. Data will need to flow across the grid using a combination of Mobile IP [10] and Ad-Hoc routing protocols such as Dynamic Source Routing Protocol (DSR)[2] [11] and Ad hoc On-Demand Distance Vector Routing (AODV)[3] [12].

Discovery Semantics and Protocols: Service description protocols are needed to describe the services provided by various components of the wireless grid. Once the services are published, a discovery protocol is needed to map the mobile resources to the services. The notion of grid service [2] can be extended to the wireless grids. Some work has been performed towards providing naming service for MANET systems [13]. The mobile nature of the wireless grid components makes it challenging to provide for discovery mechanisms across virtual organizations.

---

[2] http://www.ietf.org/internet-drafts/draft-ietf-manet-dsr-09.txt
[3] http://www.ietf.org/internet-drafts/draft-ietf-manet-aodv-13.txt

Security: Because of the inherent nature of the wireless connection, the diversity of the link quality, the potential unreliability of the end-devices, the power constraints of the mobile device, and the enforcement of security and privacy policies all present major challenges in the wireless grid environment. Effective security requires adequate computational power to execute the security algorithms in acceptable times. In addition, sufficient radio power is required to achieve an effective signal-to-noise ratio (in the face of encrypted signaling streams) and to close the link. This suggests a careful husbanding of access points and the hand-over to ensure that the minimum possible power is required from each of the wireless devices.

Policy Management: Since the end-devices or nodes can be power constrained, one cannot assume that the devices are capable of running complex protocols such as Lightweight Directory Access Protocol (LDAP) or Common Open Policy Service (COPS). Grid architecture designers need to address policies that govern the usage, privileges, access to resources, sharing level agreements, quality of service, and the composability and the automated resolution of contradictory policies among organizations; as well as other technical issues mentioned above.

6 CONCLUSIONS

Wireless grids can take ubiquitous computing to the next level by providing seamless wireless extensions to the wired grid. Mobile devices can share resources and overcome their power limitation using the grid architecture. Possible grid layouts include homogenous grid, intra-grid and inter-grid. These can be deployed to provide data, computing and utility services.

These layouts can be handled using a grid topology, which includes a combination of nodes, Root Stations and Base Stations. The topology and the architecture should address various technical challenges such as configuration, routing, discovery, security and policy management. This will dictate the ultimate ability of the wireless grids to meet the requirements at the enterprise, partner, and service levels.

ACKNOWLEDGMENTS

We thank Hongfei Tiang, Szei Hwei Ong and Madhav Srimadh for their valuable assistance.

REFERENCES

[1] I. Foster, “What is the Grid? A Three Point Checklist,” Argonne National Laboratory, http://www- fp.mcs.anl.gov/~foster/Articles/WhatIsTheGrid.pdf, 2002.

[2] I. Foster, C. Kesselman, S. Tuecke, “The Anatomy of the Grid: Enabling Scalable Virtual Organizations,” International J. Supercomputer Applications, 15(3), 2001.

[3] W. Gentzsch, “Grid Computing: A New Technology for the Advanced Web,” White Paper, Sun Microsystems, Inc., Palo Alto, CA, 2001.

[4] S. H. Ong, “Grid Computing: Business Policy and Implications,” Master’s Thesis, MIT, Cambridge, MA, 2003.

[5] H. Tiang, “Grid Computing as an Integrating Force in Virtual Enterprises,” Master’s Thesis, MIT, Cambridge, MA, 2003.

[6] A-L Barabasi, V. W. Freeh, H. Jeong, J. B. Brockman, “Parasitic Computing,” Nature, Vol. 412, 30 August 2001.

[7] S. Nesargi, R. Prakash, “MANETconf: Configuration of Hosts in a Mobile Ad Hoc Network,” in Proc. of INFOCOM’02, 1059-1068, 2002.

[8] N. H. Vaidya, “Weak Duplicate Address Detection in Mobile Ad Hoc Networks,” MIBIHOC2002, June 2002.

[9] M. Mohsin, R. Prakash, “IP Address Assignment in a Mobile Ad Hoc Network,” IEEE Military Communications Conference (MILCOM 2002), Vol. 2, 856-861, October 2002.

[10] C. Perkins, “IP Mobility Support,” Request for Comments (Proposed Standard) 2002, Internet Engineering Task Force, October 1996.

[11] D. B. Johnson, D. A. Maltz, “Dynamic Source Routing in Ad-Hoc Wireless Networks,” T. Imielinski and H. Korth, editors, Mobile Computing, 1996, Kluwer Academic Publishers.

[12] C. Perkins, E. Royer, “Ad Hoc On-Demand Distance Vector Routing,” in 2nd IEEE Workshop on Selected Areas in Communication, 90–100, February 1999.

[13] W. Adjie-Winoto, E.Schwartz, H. Balakrishnan, J. Lilley, “The Design and Implementation of an Intentional Naming System,” Proc. 17th ACM SOSP, Kiawah Island, SC, Dec. 1999.